DATA PROCESSING DIGEST

1140 South Robertson Blvd. Los Angeles 35, California

a service of Canning, Sisson and Associates

VOLUME 4 NUMBER 4

APRIL, 1958

General Information

AUTOMATION IN CHECK HANDLING

BURROUGHS CLEARING HOUSE, February 1958; pages 35-39, 104-108.

The recommendations made by the American Bankers Association for the design of checks for electronic banking systems are reviewed by the Editors ((see also DPD: July 1957, page 9)). The article then continues with suggestions for intermediate steps to be taken by banks in preparing for automatic check-handling. Banks should begin studying their own requirements for account-numbering systems, and should begin plans for imprinting both names and account numbers on all checks. It is estimated that it will take from 18 months to two years before "the pipelines are sufficiently filled with numbered checks and deposits so that a bank derives full advantage of the changeover [to automatic machine handling methods]."

Meanwhile, the Editors list some guideposts which banks should follow. In the matter of designing account numbers: "The account number... should have a minimum number of digits and provide the greatest amount of flexibility possible to be effective for future growth and changes under an automated system.... The number should be easily assigned, readily adaptable to item sorting, and should facilitate file reference work. It should indicate the type of account, controls, alphabetic breaks, and branch designation.... should be self-checking... and compatible with the characteristics of the equipment in use."

Three basic systems, along with their advantages and disadvantages, are described. These are: the numeric, alpha-numeric, and modified alpha-numeric systems. The most flexible of these is the last, in which each alphabetic group is divided into numbered subgroups, and accounts are numbered consecutively within a group. An adequate break-quantity is allowed between groups to allow for anticipated business expansion over a number of years.

The printing of checks to conform with the ABA specifications introduces some problems. It is estimated that 14 billion checks will be used annually by 1960. The added burden on the check-printing industry imposed by the imprinting of checks and deposit tickets may

CONTENTS

Check-printing problems

have to be transferred in part to "on-premises" equipment in the banks. In addition, new quality controls will have to be instituted to increase the accuracy of printing, cutting, and binding which pre-qualifying requires.

Customer relations is another area demanding the attention of bankers in the changeover to automatic methods. The design and layout of customers' checks and accounting records may have to be changed to allow for the coding line. Revisions may mean new printing plates, which raises the question of who will pay for such charges.

It is not just the large banks which will need to convert to account numbering systems. Small banks will be persuaded to do so by the Federal Reserve banks for the full use of electronic check-handling equipment. However, small banks may find some values in account-numbering even without the electronic equipment. For example, by using an account number and their conventional bookkeeping machines, these banks can adopt a pre-audit plan on personal and special checking accounts, using the account number to prove posting to the correct account for the first item and to prove the accuracy of the old balance picked up. This eliminates one posting run and one extra handling of all items. Moreover, centralized data processing centers may be established for small banks on an area basis.

The article ends with a step-by-step description of a typical computer operation in the checking account application.

BRITISH "ERA" CALLED MISSING LINK TO BANK AUTOMATION

JOURNAL OF MACHINE ACCOUNTING, December 1957; page 29.

The ERA is a character recognition device produced by Solartron Electronic Group, Ltd., of Surrey, England. It reads both alphabetical and numerical information with a photoelectric reader. It is being used presently by a large English retail chain store to read information from cash register slips sent to the store's central office by its branches. The machine is planned as a part of a building block series to cover many inputs, such as cash registers, microfilm records, checks, invoices, time clock cards, etc. Punched cards or magnetic tape output may be produced for computer use. The manufacturer sees the ERA as a solution to the problems of banks in eliminating the use of magnetic ink. The machine is expected to sell in the United States for $25,000.

BANK REDUCES ACCOUNT POSTING TO SINGLE OPERATION

OFFICE MANAGEMENT, March 1958; pages 28-30, 72-76.

The New York Trust Company has installed five NCR Postronics to handle "on-us" check posting. Future plans include the addition of

other NCR equipment designed for use with the Postronic. These include a magnetic character sorter, a magnetic character imprinter, a lister, and a punched paper tape unit. Checking accounts are given numbers which include an identification digit derived from the account number by a mathematical formula, an example of which is included. This is used by the machine to match the account number on the ledger sheet with the account number being posted. If these do not match, the machine will not accept the remainder of the posting operation.

THINKING AHEAD: HOW NEAR IS THE AUTOMATIC OFFICE?

Robert E. Slater, John Hancock Mutual Life Insurance Co.
HARVARD BUSINESS REVIEW, March-April 1958; pages 27-36, 160-176.

The "automatic office" is not yet a standardized entity, and in fact, is showing signs of changing not only its own characteristics, but those of the business organization it serves. Quality control is going to become increasingly important. Accuracy in input information is absolutely necessary because the consequences of errors are more severe in the electronic system than in a manual system. "If they interfere with machine operation or necessitate correction of basic records or repetition of machine runs, the resulting expense is very substantial indeed."

Computer specialist will gain organization power

The organization structure will tend to become less departmentalized, with departments evolving into "staff organizations planning the various operations instead of carrying them out." With fewer functioning divisions, there will be fewer administrative jobs. The man who directs the data-processing operation will gain more power; formerly independent functions will become dependent upon his operations, causing his growth to look "suspiciously like empire building to those who, in effect, will relinquish operational control over areas of the business they formerly ran." Top management will have to rethink key executive responsibilities early in the planning to make the adjustments gradually.

Contrary to most thinking, this author has found that it is not necessary to use only people with extensive company experience in the electronic data processing function. In fact, he believes that "companies are going to be less and less willing to train personnel to control their computers," and they will look to educational institutions to take over most of this responsibility. Moreover, there will begin to develop a group of computer specialists who will be concerned with how information is processed rather than why. While their ingenuity in the operation of the computer will be a great advantage, they will not have any great loyalty to the company and will not respond to traditional company orientation and communication programs. These systems specialists are not likely to be sources either for ideas or future managers as departmental people have been. Thus, "one source of potential future managers will dry up because there will be no jobs for the apprentices to do while they learn the business."

AN INTRODUCTION TO AUTOMATIC COMPUTERS

Ned Chapin, Stanford Research Institute, Menlo Park, California
Published by D. Van Nostrand Company, 1957. $8.75.

This is a careful and thorough job of computer documentation. It is a newly published, revised, and enlarged edition of the author's book published in 1955 and reviewed in DPD, December 1955, page 8.

Total EDP program planning

The chapter bibliographies, suggestions for further study, glossary, and extensive referencing suggest its use as a text or reference book in a number of related, but not necessarily interdependent courses in EDP. An instructor could select those sections which would apply to his particular course and assign reading and study from them, covering the material briefly or in detail, depending upon the emphasis of the course. Because of its copious information, the book is not easily read in the "popular" sense.

The contents may be divided into three phases of a total EDP program (although they are not conveniently labelled as such): planning for use of a computer in a business; systems design and programing for the computer; and understanding the operation of the computer.

Of the 423 pages, more than half are given to portions of the latter two phases, beginning with the site preparation and personnel selection. Within these portions, probably the most detailed explanations are those on the logical operations and the programing of the computer. A rather complete exposition is given, also, of a method for justifying the purchase of a computer (phase 1). One section is devoted to the important problem of sorting, a subject which was not covered in the original edition.

PROCEEDINGS, CONFERENCE OF AMERICAN SOCIETY FOR QUALITY CONTROL

The proceedings of the Administrative Applications Division of ASQC Conference held February, 1958, are published, and contain 12 papers, covering various aspects of management responsibility in operations research or decision-making. These include a paper on the Cummins Engine Company plan for integrated process control, an application reviewed in DPD November 1957, page 3. Persons interested in this excellent case study, who were unable to obtain the publication reviewed previously, may wish to get these proceedings. Another paper of interest which was presented at the Conference, but not included in the published proceedings is "Statistical Methods in the Clerical Field," by W. E. Jones. This suggests methods for collecting data on errors, for compiling the statistics on error rate, and for removing the causes of the errors. A bibliography and typical examples of clerical quality control are included. Write to American Society for Quality Control, Inc., Room 6197, Plankinton Building, 161 West Wisconsin Avenue, Milwaukee, Wisconsin, for both the proceedings and the paper.

PROGRESS REPORT IN CHEMICAL LITERATURE RETRIEVAL

Interscience Publishers, New York, 1957. $4.75.

This is Volume I of a series called "Advances in Documentation and Library Science," edited by Dr. Jesse H. Shera of Western Reserve School of Library Science. The fields will cover science, law, administration, medicine, and others, and will include reviews of current research, conference proceedings, monographs or articles too long for inclusion in periodicals, and other written material of value to persons concerned with the problems of documentation.

This first volume, edited by Gilbert L. Peakes of the Bakelite Company, and Allen Kent and James W. Perry of Western Reserve University, contains the proceedings of two symposiums sponsored by the Division of Chemical Literature of the American Chemical Society, a group which has been active in the solution of information retrieval problems. The papers include methods of coding information, as well as manual and mechanical methods of retrieval, not only in chemical fields, but also medicine and general scientific areas.

THE IMPACT OF COMPUTERS ON MULTIPLE-LINE OPERATIONS

John B. DeNault, Farmers Insurance Group, Los Angeles
THE CONTROLLER, March 1958; pages 116, 117, 143, 144.

The use of the IBM 705 at Farmers Insurance has been the subject of several articles listed in DPD ((see May 1957, page 11, and December 1957, page 13)). This article covers the same ground, relating the conversion of the punched card system to the 705 with magnetic tape files.

Systems Engineering

NEW ORGANIZATION ANALYSIS TECHNIQUES

Irwin P. Lazarus, American Colortype Company, Chicago
JOURNAL OF INDUSTRIAL ENGINEERING, January-February 1958; pages 23-27.

Five types of graphic representations of business organizations are described. The first of these is the Linear Responsibility Chart, which is a graphic picture of the Organization Manual. "Across the top of the chart are listed the segments of the organization that one is interested in analyzing. Up the left side of the chart are listed the functions that are under analysis. In the body of the chart are drawn graphic symbols to depict the amount of authority and responsibility that each organizational segment exercises over the functions in question. These eight graphic symbols vary from the actual doing of the work to calling a person in for an exchange of views."

Chart shows function responsibilities

The second type is the Multiple Correlation Chart, which is similar in layout to the first type, but easier to read. Instead of using symbols to represent the authority and responsibility, "the Multiple Correlation Chart is made by cutting from the source document the actual sentences or paragraphs that describe the amount of authority and responsibility that each organizational segment exercises over the function under investigation.... These source documents are: 1. The Organizational Manual, 2. Job Descriptions, 3. Standard Procedures and Method Descriptions, 4. Policy Statements, 5. Actual Operations."

When some of the source documents are missing, a substitute may be found by using the Spectrum Analysis method. In this, the chart is filled in with whatever information is available from policy statements or various regulations. "After interviewing the personnel of the organizational segment described on the card, the actual method of performance was determined. This was then shown by attaching colored cellophane directly over the printed matter; with the color of the cellophane coded to represent the method of performance.... This chart clearly indicated deficiencies in the performance of the organization and was highly effective in gaining top management attention."

Organization chart reveals real lines of authority

The fourth type is the informal organization chart which has the same form as the formal organization chart of the company except that instead of the usual line of authority descending from the top of the chart down, lines are drawn to show the actual route of reporting within the organization. "Dash lines signify formal lines of supposed authority and responsibility. Solid lines signify actual lines of habitual reporting and directing." In the example given, the chart shows individuals reporting two levels above rather than to their immediate superior. Such a chart directs the attention of management to the need for training and some kind of conferences to "reconcile misunderstandings in the flow of information and determine the optimum route of reporting."

Confused lines of authority show up

The fifth technique is that of "having all key individuals in the organization draw the organization chart as they picture it. This technique is extremely useful in companies where no organization chart exists and all individuals are confused because of the lack of a chart. Following the collection of all key people's drawings, a reconciliation is made and the optimum chart decided and described to all people concerned." In the example given, one chart shows the organization as the Vice-President and General Manager of the company visualizes it. He has six organizational segments reporting to the Chief Engineer. The Chief Engineer, in his own chart, shows only four, one of which is not mentioned by the General Manager. The Chief Engineer shows the Treasurer as a staff officer with the General Manager sometimes reporting to him. The General Manager indicates the Treasurer is a line official directly between himself and the President. However, the Treasurer ("a rather modest individual"), depicts himself as a staff official reporting usually to the General Manager and sometimes to the President. Eventually the charts were reconciled, and the article shows the solution.

AUTOMATION FOR SMALL-LOT PRODUCERS

AUTOMATION, March 1958, pages 54-68.

A Special Report is given by four members of General Electric's Automation Task Force, which has been devoting full time to "formulate a basic approach to the problems of automation and mechanization... to perform operations analyses... to recommend automation projects to management...."

The initial study of the manufacturing activity (Light Military Electronic Equipment) showed that the "low-quantity production of a wide diversity of highly complex electronic equipment.... makes for a fairly high engineering content... so that automation in the engineering section is very desirable in order to increase productivity." It was found that about 80 per cent of the data-handling in the Department could be made automatic. The study also revealed "common denominators among the variety of units produced," suggesting "the possibility of creating volume where none had appeared to exist."

Each operation was studied and recorded on work sheets under the classifications "automatable" or "nonautomatable." The later was identified using these criteria: 1. Does it require a human decision? 2. Is it outside the scope of direct control by the Department? 3. Is the necessary automation technology available? Simple formulas were then used to select the areas best suited to automation.

Five concepts of automation feasibility

The Task Force has developed five concepts which appear to be applicable to other businesses regardless of the product or market. These are:

1. Universally Applicable Systems Concept: "The most benefits can be gotten from automation when it is thought of as a universally

applicable systems philosophy including industrial research, economic analysis, product design, machine design, machine construction, and machine utilization."

2. Designing Your Business Concept: "It is both possible and desirable to reshape or mold an apparently incompatible business into a form to which automation can be applied."

3. Systems Planning Concept: "Automation can best be approached through an evolutionary, long-range program, based on systems planning."

4. Efficiency Concept: "Systems planning attendant to a long-range automation program is self-supporting because of recommendations that increase productivity and reduce costs--whether or not the ultimate goal of across-the-board automation is ever achieved."

5. Do-It-Yourself Concept: "The ideal equipment...is normally not available on the open market....Moreover, the use of a systems approach presupposes an intimate knowledge of the products and the market for the particular business. So it is best to work with a staff of permanent employees drawn from the appropriate functional sections of the plant organization."

Three additional articles discuss the automation of fabrication, quality control, and information processing.

APPLICATION OF ELECTRONIC DATA TO BILLING AND ACCOUNTING

Ralph C. McCollum, The Peoples Gas Light and Coke Co., Chicago
JOURNAL OF MACHINE ACCOUNTING, December 1957; pages 6-13, 34-39.

In designing an electronic data processing system, it is easy to overlook some details which were brought to light in the early stages, unless these points are noted specifically. It is suggested that every machine or manual operation and every tab card, magnetic tape and form be represented by a requirements sheet. ((No clear description or example is included.)) These sheets would provide the programers with a description of each operation, and would be useful in noting the effects of changes in any part of the system. Also, by checking off the completed items, a rough idea of progress can be gained. During these stages memorandums should be prepared giving reasons for each decision.

Classify the exceptions

The system should be designed on the basis of normal cases, with the handling of exceptions injected into the system later. Normal cases include the "going concern" procedures, the record changing procedures, the inquiry handling procedures, and the conversion procedures. Exceptions should be studied to find the reason behind each one, to classify them into types, such as permanent, temporary, and combinations of several conditions; and to determine the undesirable effects that result from their existence.

Exceptions require additional checking operations on input data and additional testing and selection operations throughout the procedures. They increase the information capacity requirements of the system, and the total volume of instructions needed.

The need for ultimate changes in the system should be anticipated in five areas: 1. Eliminating duplicate records; 2. Mechanizing input data and developing information codes; 3. Mechanizing areas of personal judgment; 4. Eliminating manual maintenance of codes and standards; 5. Developing new kinds of reports.

Examples are given to illustrate each of these areas of change. For instance, in mechanizing the judgment area, The Peoples did extensive research on the causes of high bill complaints. The result of the research was a discovery that the same formula could be used for averaging bills for non-heating accounts that is used for heating accounts, thus simplifying the judgment factor in the customer billing operation.

Management Decision-making Techniques

BUSINESS GAMES — PLAY ONE!

G. R. Andlinger, McKinsey & Company, Inc.
HARVARD BUSINESS REVIEW, March-April 1958; pages 115-125.

Game-playing without a computer

A business game is described which can be played without the aid of an electronic computer. The game was developed by McKinsey & Company. It is similar to the game developed by the American Management Association, except that it deals with capital goods companies rather than consumer goods companies, and it introduces a realistic element of time lags between decision making and results.

The Business Management Game "incorporates the elements of a one-product capital goods company which is competing with one or two other companies for the same market.... The Market, Marketing, Advertising, Research and Development, Production, Finance, and Competition." The game may be played by two or three teams, or single individuals, placed in separate rooms. An umpire group of three or four persons who know the rules calculate the operating results of each team's set of decisions, and decide the trend of the market at the beginning of the game. They have a random number table for computing the various parameters and trends. Each team has a game board and a quantity of models or counters, representing salesmen, factories, production units, and accounts receivable sheets.

These are placed on the same board to represent the characteristics of the company in any one quarter. Decision sheets are filled in in duplicate, with the carbon given to the umpire group for computing and reporting back to the teams.

"To avoid an overly competitive atmosphere, it is helpful to have each team spell out its general objectives and its management philosophy at the beginning of play and then elicit each team's own evaluation of the soundness of its policy framework at a postgame critique session."

It takes about seven years of operation (at about 15–20 minutes per quarter) before mature game planning, and tough competitive battles begin.

"...a generalized business game such as the Business Management Game must be considered only a beginning....In time we should be able to develop specific games...that simulate the functioning of one specific company and its major competitors. The training value of such games and the amount of insight and skill which can be drived from them may be a genuine 'breakthrough' in executive development."

Sets of the game with a reprint of the article may be ordered from #A1, Readers Service Department, Harvard Business Review, Boston 63, Massachusetts. Price, $1.00.

AN INVESTIGATION OF SOME QUANTITATIVE RELATIONSHIPS BETWEEN BREAK-EVEN POINT ANALYSIS AND ECONOMIC LOT-SIZE THEORY

Wayland P. Smith, Michigan State University
JOURNAL OF INDUSTRIAL ENGINEERING, January-February 1958, pages 52-57.

"Two common tools utilized to evaluate the economic potential of alternative ways of performing a specified task are (A) Break-Even Point Analysis, and (B) Economic Lot-Size Theory."

"In retrospect one wonders why these two schemes have never been combined into one common model. On the other hand, they have primarily been developed to solve two problems that at least on the surface appear to be quite divergent. The purpose of this paper is to show that these problems are not divergent, that they have much in common, and that they are actually very much interrelated." A not-too-complex mathematical formula is developed that relates the two theories.

To explore a practical case, the paper examines a problem "involving three alternatives: namely, the age-old problem of whether it is better to use an engine lathe, a turret lathe, or an automatic lathe to perform a specified manufacturing operation."

In the example given, for instance, three break-even points are developed. "The engine lathe is the most economical method when

the quantity required is greater than 0 and less than 290 pieces. The turret lathe is the most economical method when the quantity required is greater than 290 pieces and less than 800 pieces. The automatic lathe is the most economical when the quantity required is in excess of 800 pieces."

Conclusions reached are:

Break-even model can produce errors

1. "The use of the simple break-even model to decide between several alternative methods of performing a task can give rise to serious errors.

2. These errors are particularly pronounced when--

   a. The production rates are considerably higher than the demand rates.

   b. The storage and inventory costs are significantly large.

   c. There is a substantial difference between the ratio of set-up cost and storage cost for the several alternative methods.

3. The combined model is not significantly more difficult to work with when modern computing techniques and equipment are considered."

Applications

ILLINOIS DRIVERS LICENSE DIVISION

The Drivers License Division of Illinois State has installed an IBM 650 Tape computer to process all transactions concerning drivers' licenses. A master file contains all the records formerly kept in 13 separate files. All the actions resulting from a driver's new record are recorded on another tape which is used to print the necessary warning letters, suspensions, revocations, cancellations and new abstract of records. Each revocation or suspension order is accompanied by an abstract of the driver's record for review by driver control officers before final issue. The system has made possible the adoption of a new type of permanent driver license number which establishes a positive identification, and eliminates confusion of records of drivers with identical names. The number becomes an individual control number for use throughout all future record keeping and processing operations.

OTAC'S BIZMAC

COMPUTING NEWS, March 1, 1958; pages 3-5.

OTAC (Army Ordnance Tank-Automotive Command) at Detroit uses an RCA Bizmac system in processing an inventory of 200,000 kinds of parts for a million vehicles in all parts of the world.

The installation was completed in October, 1956. About a third of the projected program is in operation. The complete system will have a three-fold job: keep an inventory on spare parts, catalog parts and prices for rapid manual preparation, and forecast supply requirements by 12, 24, and 36 months and project budget costs.

The system consists of 220 units weighing a total of 217 tons, and requiring 250 tons of air conditioning. Personnel consists of seven engineers, 27 technicians, 8 analysts, 10 senior and 9 junior programers, and 45 operators, schedulers and other staff members. All personnel were chosen because of their familiarity with the OTAC supply problem. A continuous training program is maintained at the Computation Laboratory of Wayne State University.

TEXTILE INDUSTRY INSTALLS RAMAC

JOURNAL OF MACHINE ACCOUNTING, December 1957; pages 29, 30.

The United Merchants and Manufacturers, Inc. is using the IBM 305 RAMAC experimentally in solving problems of merchandising control in its textile converting operations. Sales activities in textiles in all parts of the country are processed to reflect the buying trend. This information is used to adjust production and marketing plans in quicker response to the fashion changes than has been possible.

Programing

COMPUTER SIMULATION SYSTEMS — MACHINE LANGUAGE

COMMUNICATIONS OF A.C.M., February 1958; page 4.

Some programs have been written for computers which simulate the operation of other computers. The advantage of such a procedure is the use of an available computer for runs which cannot be handled by the computer for which they were intended because of overloading or a non-operating condition. A table is given which lists 13 such programs with the organizations which produced them. On page 1 is an announcement of such a program prepared by The Council for Economic and Industry Research for their 704, simulating the operation of a 650.

Equipment

NEW COMPUTERS — A REPORT FROM THE MANUFACTURERS

Proceedings of Symposium sponsored by L. A. Chapter of A. C. M.

The Proceedings of the Symposium on New Computers, sponsored by the Los Angeles Chapter of the Association for Computing Machinery March 1, 1957, contains nine papers on new large- and medium-scale computers. These include: NCR 304, ElectroData Cardatron and DataFile, DATAmatic 1000, Bizmac II, X308 (Univac), IBM 709, IBM STRETCH, Philco S-2000, and Alwac 800. The proceedings may be ordered from the Association for Computing Machinery, 2 East 63rd Street, New York 21, New York.

SELECTING AREAS IN PUNCHED PAPER TAPE ACCORDING TO ADDRESS

COMPUTERS AND AUTOMATION, February 1958; pages 17, 18.

Friden has placed on the market the Selectadata, a device for automatic selection and sorting of data encoded in punched tape. The device eliminates the need for converting tape to punched cards in order to select the desired data. A row of seven selection keys on the front panel offers a choice of up to 127 different address codes.

Comment

SORTING

SORT--To arrange items of information according to rules dependent upon a key or field contained by the items. (Reference 1, pg. 20.)

Sorting is one of the less glamorous and yet more important data processing functions. There is a tendency to disregard sorting during systems design. But the choice of effective sorting methods can be a vital, cost-saving, and challenging study. Sorting on the computer cannot be avoided except in the more trivial EDP applications; one must eventually face it. And if it must be done, it pays to try to use sort programs which are the most efficient for the application. As an EDP installation becomes fully loaded, a reduction in sort times can permit the addition of applications which contribute to better management control and so to company profits.

Four systems design decisions

The four basic decisions which the systems designer has to make in regard to sorting are discussed below.

1. Is sorting needed? Only a complete random access system will eliminate sorting and often even with random-access some sorting is desirable to reduce costs or to provide control data. For any magnetic tape system many sorts will be needed: to put transaction data into file sequence, output data into report sequence, and between runs to re-sequence data. (Ref. 2, pg. 134.)

2. Should the sorting be done on the computer, on an off-line electronic unit (available now only with UNIVAC File Computer Model 1) or on punched-card sorters? Where data is already in card form, as at input or just before printing on a tabulator, card sorting may be the most economical method. But when the cost of operators and control is included, sorting is often cheaper on the modern data-processing systems, if well-designed sort programs are used. Sorts between computer runs must be handled on tape, as card punching and reading costs are prohibitive.

3. What method of computer sorting should be used? The answer to this question depends upon the computer and the data, but some basic techniques are mentioned below.

The basic factors in sorting are those describing the data--

- Number of computer words or characters per item
- Number of digits in the key (or identifying field) upon which the sort is to be made
- Number of items to be sorted

and those describing the quipment--

- Number of tape units available

- Speed of transfer of data between tape and computer

- Internal processing speed

- Size of internal memory. (Ref. 10, pg. 2.)

Given a sorting method and these factors, sorting time, and therefore cost, can be determined. (References 2, 10.)

Computer sorting methods

In nearly all cases the most economical method of sorting with magnetic tape systems is sorting by merging. In its simplest form, the input data is divided into two groups which are merged and remerged. Each merge, often called a "pass," results in longer and longer sequences until finally the items are all in sequence. (Ref. 11.)

The merge method can be effectively speeded up in several ways:

a. The first few passes can be replaced by internal computer processing. The internal sort results in sequenced groups, which are then sorted by tape merges. The speed of processing is materially affected by the program for the internal sort, which, therefore, deserves careful study. (Ref. 2, 10.) The larger the internal merged sequences, and the faster the computer, the faster the sorting. Special commands often assist in programing sorts. (Ref. 2, pg. 137.)

b. Careful attention to use of input and output buffers permit faster transfer of data to and from tape during the merging passes. (Ref. 2, pg. 138.)

c. Use of several tapes for input and output (often 5 in and 5 out) permit the merging or meshing of more data each pass. (Ref. 2, pg. 155; ref. 10, pg. 15.)

d. During merging, additional comparisons can be made to permit the building of sequences faster (e.g., comparing present items with item just read out as well as with each other).

Good programers can often effect savings of 10-30% in sorting time by attention to these points.

The other major type of sorting is radix sorting which is exemplified by the operation of a card sorter--the column-by-column sort. There are various sophisticated forms of radix sorting in which the number base may vary. (Ref. 2, pg. 155; ref. 10, pg. 5.) A degenerate case is called sorting by distribution, wherein the data is distributed among many pockets and then collected in the proper se-

quence. (Ref. 2, pg. 152; ref. 3; ref. 9; ref. 10, pg. 26.) This method is often proposed for sorting in a random-access unit.

Routines furnished by manufacturer

4. The final decision to be made in regard to sorting is whether to use a routine provided by the manufacturer (Ref. 4, 5), or one developed by your own programers. Usually, at the beginning of an operation the manufacturers' programs are used, since programing time is in short supply. Later, as computer capacity becomes the item in short supply, most groups find that programing their own sorts is worthwhile. (Ref. 6, 7, 8.) It is almost always possible to write a program which will sort your specific data faster than will the manufacturer's routine, which must be written to cover many situations.

Finally, we would like to point out some of the difficulties encountered when sorting on a computer.

- It is very difficult to sort data in one run where the items are of various lengths, especially if complex internal sort programs are used. Either the items of a given length must be sorted separately and then merged; or they must be sorted as if they were all as long as the longest. The method chosen will be the one that minimizes total processing time.

- Sometimes a key field is split and is located in several parts of the item (Ref. 7); this complicates the program, and can make off-line sorting, if used, difficult. If the items of the sort contain various transactions with keys located in different fields, a preliminary edit to standardize the fields will almost certainly be necessary.

- It may be difficult to check the sorting. It is desirable at least to take and check a count of the number of items of each pass; and often one would like to take and check control totals. Often, either of these checks are impossible on off-line sorts, and need special attention even in computer sorts. Any sort, especially off-line sorting, should be followed by a sequence check in the first computer program following the sort.

- If the computer fails in the middle of a sort it is undesirable to have to rerun it completely. Thus, restart data and procedures should be considered for inclusion in the program at appropriate intervals.

To summarize, it is generally worthwhile to program sort routines carefully, so that the computer time (and therefore money) can be saved.

REFERENCES

1. First Glossary of Computer Terminology, June 1954; Association for Computing Machinery, 2 East 63rd St., New York 21, N.Y.

2. Edward H. Friend, "Sorting on Electronic Computer Systems," Journal of the Association for Computing Machinery, Vol. 3, No. 3, July 1956; pages 134-168.

3. Ida Rhodes and Mary Stevens, "Preliminary Report on a Combined Sorting-File Merging Method for Electronic Data Processing," National Bureau of Standards Report 3155, U. S. Dept. of Commerce, Washington, D. C.

4. "702 Sorting Techniques," EDPM Program Brief 8, Form 22-6656-0, IBM Corp., 590 Madison Ave., New York 22, N. Y.

5. "IBM 705 Generalized Sort Program--Sort 57," Form 32-7880, IBM Corp. (address above).

6. Arthur Shapiro, "Sequence Check and Sort," The Programmer, Vol. 2, No. 6, Nov.-Dec. 1955, Remington Rand Corp., 315 Fourth Ave., New York 10, N.Y.

7. Martin Goetz, "4-Way-20-Word Split Key Item Sort Run," The Programmer, Vol. 3, No. 1, Jan.-Feb. 1956 (see address above).

8. Dave Savidge, "The Arranger Routine--Ten-Word Item--10x10." The Programmer, Vol. 2, No. 5, Sept.-Oct. 1955 (see address above).

9. E. J. Isaac and R. C. Singleton, "Sorting by Address Calculation," Journal of the Association for Computing Machinery, Vol. 3, No. 2, July 1956, pages 169-174 (see address above).

10. Marc Shiowitz, "Digital Computer Programs for Data-Sorting Problems," lecture notes, Summer Session, Wayne University, Detroit, 1954 (out of print; a revised edition may be published. Write us if you are interested).

11. Richard G. Canning, "How the Four-Tape Sorter Simplifies Storage," Control Engineering, February 1957, pages 95-97.

References

The addresses of publishers and periodicals mentioned in this issue of DATA PROCESSING DIGEST are listed below for your convenience in obtaining further information about the articles or books listed.

Automation
Penton Publishing Company
Penton Building
Cleveland 13, Ohio

Burroughs Clearing House
2nd and Burroughs Avenue
Detroit 32, Michigan

Communications of A. C. M.
5800 N. Marvine Street
Philadelphia 41, Pennsylvania

Computers and Automation
815 Washington Street
Newtonville 60, Massachusetts

Computing News
Jackson Granholm
12805 - 64th Avenue South
Seattle 88, Washington

The Controller
Two Park Avenue
New York 16, New York

Harvard Business Review
Soldiers Field Station
Boston 63, Massachusetts

Interscience Publishers, Inc.
250 Fifth Avenue
New York 1, New York

Journal of Industrial Engineering
American Institute of Industrial Engineers
145 North High Street
Columbus 15, Ohio

Journal of Machine Accounting
208 South Main Street
Paris, Illinois

Office Management
212 Fifth Avenue
New York 10, New York

DATA PROCESSING DIGEST is published each month by Canning, Sisson and Associates, 1140 South Robertson Blvd., Los Angeles 35, California. Subscription rate: $24.00 per year. Foreign postage (exclusive of Canada and Mexico): $2.50 additional. Single copies: $3.00 when available. Editor: Margaret Milligan.

Training

Linear Programming--One-week Basic Course, sponsored by Marquette University Management Center and The American Institute of Industrial Engineers

Date: April 28-May 2, 1958
Place: Milwaukee, Wisconsin (Marquette University)
Fee: $125
Information: Management Center, Registration Clerk, Marquette University, Milwaukee 3, Wisconsin

"The Company President Looks at Electronic Data Processing," sponsored by the University of California at Los Angeles

Date: May 1, 2, 1958
Place: Lake Arrowhead, California (UCLA Residential Conference Center)
For whom: Company presidents and vice presidents
Fee: $100 (includes room and board)
Information: Mr. Sam Houston, Room 3116, Engineering Building, University of California, Los Angeles 24, California

Operations Research in Production and Inventory Control, sponsored by Case Institute of Technology

Date: June 2-13, 1958
Place: Case Institute, Cleveland, Ohio
Information: R. L. Bell, Engineering Administration Department, Case Institute of Technology, 10900 Euclid Avenue, Cleveland, Ohio

Summer Engineering Seminars, sponsored by The Pennsylvania State University
Introduction to Computer Programming, June 16 to 21
Automatic Data Processing in Business and Industry, June 23 to 28
Scientific and Engineering Computation, July 13 to 25
Mathematical Methods for Management, August 3 to 8
For information, write: Extension Conference Center, The Pennsylvania State University, University Park, Pennsylvania

Computer and Management Science Program, University of Michigan
Summer Session, 1958. Courses on computer engineering, operations research, management sciences, and business data processing are included. Session begins June 16, 1958. For information, write to: Coordinator of Engineering Summer Conferences, 2038 East Engineering Building, Ann Arbor, Michigan.

Work Simplification and Measurement Conference, sponsored by Wharton School (University of Pennsylvania)

Date: June 16-27, 1958
Place: Philadelphia, Pennsylvania (University of Pennsylvania)
Fee: $375
Information: Dr. Adrian M. McDonough, Director, Wharton Refresher Conference, Wharton School of Finance and Commerce, University of Pennsylvania, Philadelphia 4, Pa.

Advanced Business Systems Conference and Workshop, sponsored by Wharton School (University of Pennsylvania)

Date: July 7-18, 1958
Place: Philadelphia, Pennsylvania (University of Pennsylvania)
Fee: $375
Information: Same as above

Electronic Data Processing for Business and Industry (Course 10), sponsored by Canning, Sisson and Associates

Date: July 28 through August 1, 1958
Place: New York (Hotel Biltmore)
Fee: $250
Program: Emphasis on the applications aspect of electronic data processing, planning for an EDP system
For whom: Management personnel charged with setting up an EDP system
Information: Canning, Sisson and Associates, 1140 South Robertson Blvd., Los Angeles 35, California

Meetings

Conference on Automation, Operations Research & Business Planning, sponsored by University of Chicago Downtown Center

Date: April 21, 22, 1958
Place: Chicago, Illinois (Morrison Hotel)
Information: University of Chicago Downtown Center, 19 LaSalle Street, Chicago 3, Illinois

Western Joint Computer Conference

Date: May 6-9, 1958
Place: Los Angeles, California (Ambassador Hotel)
Theme: Contrasts in Computers. The last day will be devoted to Reports from the Manufacturers on small automatic computers and input/output equipment.
Information: Dr. Willis H. Ware, General Chairman, care of Rand Corp. 1700 Main Street, Santa Monica, California

Third Annual Electronic Data Processing Conference, sponsored by University of Alabama and local chapters of professional societies

Date: May 8, 9, 1958
Place: Campus of University of Alabama
Program: Papers on industrial and business uses of electronic data processing
Information: Gordon E. P. Wright, Director, Commerce Extension Services, P.O. Box 2987, University, Alabama

Operations Research Society of America (Annual Meeting)

Date: May 15, 16, 1958
Place: Boston, Massachusetts (Sheraton-Plaza Hotel)

NOMA Office Show and Conference

Date: May 25-28, 1958
Place: Chicago (Conrad Hilton Hotel)
Information: The National Office Management Association, Willow Grove, Pennsylvania

Canadian Conference for Computing and Data Processing, sponsored by University of Toronto

Date: June 9, 10, 1958
Place: Toronto, Canada (University of Toronto)
Program: The present status of the use of EDP systems in Canada
Fee: $35
Information: Mr. H. W. Rowlands, 15 Wellington Street W., Toronto, Ontario, Canada

International Automation Exposition and Congress

Date: June 9-13, 1958
Place: New York (Coliseum)
Information: Richard Rimbach Associates, Show Management, 845 Ridge Avenue, Pittsburgh 12, Pennsylvania

Association for Computing Machinery Annual Meeting

Date: June 11-13, 1958
Place: Urbana, Illinois (University of Illinois)

"Industrial Engineering--Gateway to Productivity," 9th Annual Conference of the American Institute of Industrial Engineers

Date: June 12-14, 1958
Place: Los Angeles (Hotel Statler)
Program: Fifty sessions on industrial engineering subjects will be held. In addition, during the conference, management game will be played as a demonstration of simulation technique in management-decision-making. An IBM 650 will be used for computing the results during the playing.
Information: Stanley Wolfberg, 1409 Thayer Avenue, Los Angeles 24, California

5th Annual Symposium on Computers and Data Processing, sponsored by Denver Research Institute

Date: July 24-25, 1958
Place: Denver, Colorado (Albany Hotel)
Information: C. A. Hedberg, Denver Research Institute, University of Denver, Denver 10, Colorado

Instrument Society of America Automation Conference

Date: September 15-19, 1958
Place: Philadelphia, Pennsylvania (Convention Hall)

Fourth Electronic Business Systems Conference, sponsored by Western Division of NMAA

Date: October 30, 31, 1958
Place: Seattle, Washington (Olympic Hotel)
Information: E. B. S. Conference, National Machine Accountants Assoc., Western Division, P.O. Box 134, Seattle 11, Washington

National Physical Laboratory Symposium and Electronic Computer Exhibition

Date: November 28-December 4, 1958
Place: London, England
Information: C. V. Wattenbach, Deputy Managing Director, Dictograph Telephones, Ltd., London, England

Eastern Joint Computer Conference

Date: December 1958
Place: Boston, Massachusetts

SHARED PROGRAMING GROUP

SHARE- Sept. 10-12, 1958, San Francisco